# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188



**1. AGENCY USE ONLY (Leave blank)**

**2. REPORT DATE**
August 1997 (12/10/97)

**3. REPORT TYPE AND DATES COVERED**
Final Technical Report, June 1996 – August 1997

**4. TITLE AND SUBTITLE**
Clearning Phased Array Radar Data

**5. FUNDING NUMBERS**
AFOSR-TR 97-0727

**6. AUTHOR(S)**
Dr. Andreas S. Weigend

**7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)**
formerly University of Colorado at Boulder, now Leonard N. Stern School of Business, NYU

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAMES(S) AND ADDRESS(ES)**
Air Force Office of Scientific Research
110 Duncan Avenue, Suite B115
Bolling AFB, DC 20332

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

19971217 011

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Distribution Unlimited

**12. DISTRIBUTION CODE**

**13. ABSTRACT (Maximum 200 words)**

Many military and civilian problems can be viewed as pattern recognition: given a set of measured inputs, the task is to predict the corresponding output. Typical examples range from image recognition and classification, to time series prediction and regression. Most modeling assumes that the inputs can be measured exactly, without noise. Building a model then means to construct (or "learn") a mapping from these inputs to the expected values of the outputs. The usually tacit assumption of noise-free inputs is violated in most real-world problems where only a noisy version of the "true" input is observed.
This research found that while it was possible for *time series problems* even if there is a lot of noise present, to use information from adjacent patterns in time, the problem could be solved for non-time series problems, such as the phase array radar data.
The effort lead to several papers. Results are presented on discrete hidden states (Hidden Markov models), and continuous hidden state (state space models). A paper on finding the true inputs using Independent Component Analysis is in preparation. A paper on evaluation methodology using the bootstrap also employs the state space approach.

**14. SUBJECT TERMS**
Machine Learning. Errors in Variables. Orthogonal Distance Regression. Total Least Squares. Nonlinear Modeling. Time Series.

**15. NUMBER OF PAGES**

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | UL |

# Clearning Phased Array Radar
## AFOSR Grant F49620-96-1-0240, Final Report
## AFOSR Program Manager: Jon Sjogren
## Principal Investigator: Andreas Weigend

Until December 1996: Assistant Professor, Computer Science Department, University of Colorado at Boulder

Since January 1997: Associate Professor, Information Systems Department, Leonard N. Stern School of Business, New York University

## 1. Objective

Observations, such as phased array radar data, contain noise, usually from several sources. The essence of modeling, and subsequent inference, is to extract the signal. The objective of this grant was to understand the strengths and limitation of a new algorithm called "clearning" (the combination of learning the model and cleaning the data), and to apply it to phased array radar data.

## 2. Results

The proposal was written with a three-year time horizon. Before applying the algorithm to phased array radar data, and comparing it to competing algorithms, the first goal was to understand what the algorithm can do, and what it cannot do. This was best done by relating it to the important research question: to what degree can we infer hidden states from observed data?

The key result is: hidden states can be inferred successfully for time series data. Time series data have the major advantage that adjacent patterns are indeed related to each other. This is not the case in standard, non-time-series pattern recognition problems.

The first progress report emphasized the important of constraints between the input variables to exist for clearning to work. In particular, it emphasized that the first steps of the project thus are to clarify what might be done, and what cannot be done in principle, as well as to relate clearning to source separation, and, in the case of time series, to state space modeling and Kalman filtering. This has been achieved: The following describes the research that my collaborators and I carried out in the last year in the context of finding ("hidden") variables (continuous, as in clearning, or discrete) that are a less noisy characterization of the systems than a snapshot of the raw observed signal.

Shi and Weigend [1] explore discrete hidden states, and show their usefulness for characterizing and predicting very noisy time series. This is an extension of hidden Markov models, very popular in the speech community, but hardly known in the prediction community. The key idea is: if there are different dynamics in different regimes of the time series, and these regimes last for a while, then rather than averaging over the submodels, a more

appropriate model is obtained by estimating both the regime, and the parameters of the sub-models.

The MATLAB code we wrote for these experiments is available upon request.

The power of hidden Markov models crucially depend on the time series nature of the problem. Clearning, in contrast, as well as the "gated experts" architecture (Weigend, Mangeas, and Srivastava 1996) do not exploit the time series structure and are thus both more broadly applicable and weaker.

Timmer and Weigend [2] show the power of modeling dynamic noise and observational noise separately. I had mentioned previously (Section 2.1 of the progress report) that noisy inputs can lead to an underestimation of the parameters. This paper explores this point further and shows that a case where the decay times of shocks are underestimated by two orders of magnitude when the distinction between observational and dynamic noise is ignored. While state space modeling is a powerful method, it crucially depends on the time series nature of the problem.

Another method, suggested in the progress report, is blind source separation, related to independent component analysis (ICA). In collaboration with Dr. Andrew Back I started to explore the usefulness of independent component analysis (ICA, also called blind source separation) to very noisy data, Japanese stock return, in comparison to principal component analysis (PCA). Preliminary results indicate that estimated independent components (ICs, also called "sources") fall into two distinct categories: (1) a small number of large transient shocks (with skewed distributions), and (2) approximately Gaussian random noise.

Finally, the revision of a third paper by LeBaron and Weigend [3] focusing on focuses on performance evaluation by re-sampling, profited from the distinction of different noise sources: the method described in [2] was applied to that time series of daily NYSE volume.

In summary, while these papers received attention at several conferences and workshops, and have been accepted by major journals, the answer to the first stage of the clearning question has, unfortunately, been largely negative. I currently do not see a way to extend the algorithm to non-time-series data as I had hoped: there simply is not enough information for the degrees of freedom of both moving the data and the model.

## 3. Publications

[1] Shanming SHI and Andreas S. WEIGEND "Taking Time Seriously: Hidden Markov Experts Applied to Financial Engineering." In: Proceedings of the IEEE/IAFE 1997 Conference on Computational Intelligence for Financial Engineering (CIFEr, New York, March 1997), pp. 244--252. Piscataway, NJ: IEEE Service Center.

http://www.stern.nyu.edu/~aweigend/Research/Papers/HiddenMarkov/

Abstract--Most traditional time series models are global models based on local time information: they assume that the state can be fully and locally (in time) characterized with a finite embedding space. Prediction then amounts to simple regression. Unfortunately, there are many situations in which simple regression is not sufficient to model the temporal structure in a time series. We here introduce an architecture that we call Hidden Markov Experts. It is based on Hidden Markov Models used in speech recognition research. By introducing the concept of hidden states, Hidden Markov experts model time dependency of time series explicitly as a first-order Markov model with transitions between these hidden states. Within each state, local models are applied to estimate the probability density, which can be linear or nonlinear depending on the situation. This paper first discusses the statistical framework and the learning algorithm of Hidden Markov experts, then applies them to daily S&P500 data and to high frequency currency exchange rate data. The Hidden Markov Experts have better profit than the linear and nonlinear global models. The volatilities of the time series can be characterized by the hidden states.

[2] Jens TIMMER and Andreas S. WEIGEND "Exploiting Local Relations as Soft Constraints to Improve Forecasting." Forthcoming in: International Journal of Neural Systems, Vol. 8 (1997).

http://www.stern.nyu.edu/~aweigend/Research/Papers/StateSpace

Abstract--In time series problems, noise can be divided into two categories: dynamic noise which drives the process, and observational noise which is added in the measurement process, but does not influence future values of the system. In this framework, empirical volatilities (the squared relative returns of prices) exhibit a significant amount of observational noise. To model and predict their time evolution adequately, we estimate state space models that explicitly include observational noise. We obtain relaxation times for shocks in the logarithm of volatility ranging from three weeks (for foreign exchange) to three to five months (for stock indices). In most cases, a two-dimensional hidden state is required to yield residuals that are consistent with white noise. We compare these results with ordinary autoregressive models (without a hidden state) and find that autoregressive models underestimate the relaxation times by about two orders of magnitude due to their ignoring the distinction between observational and dynamic noise. This new interpretation of the dynamics of volatility in terms of relaxators in a state space model carries over to stochastic volatility models and to GARCH models, and is useful for several problems in finance, including risk management and the pricing of derivative securities.

[3] Blake LeBARON and Andreas S. WEIGEND "A Bootstrap Evaluation of the Effect of Data Splitting on Financial Time Series." Forthcoming in: IEEE Transactions on Neural Networks, Vol 9 (1998).

http://www.stern.nyu.edu/~aweigend/Research/Papers/Bootstrap/

Abstract: This article exposes problems of the commonly used technique of splitting the available data into training, validation, and test sets that are held fixed, warns about drawing too strong conclusions from such static splits, and shows potential pitfalls of ignoring variability across splits. Using a bootstrap or resampling method, we compare the uncertainty in the solution stemming from the data splitting with neural network specific uncertainties (parameter initialization, choice of number of hidden units, etc.). We present two results on data from the New York Stock Exchange. First, the variation due to different resamplings is significantly larger than the variation due to different network conditions. This result implies that it is important to not over-interpret a model (or an ensemble of models) estimated on one specific split of the data. Second, on each split, the neural network solution with early stopping is very close to a linear model; no significant nonlinearities are extracted.

## 4. Presentations

Time Series Analysis and Financial Modeling Johns Hopkins (Baltimore, Jan 9, 1998)

Modeling Volatility Using State Space Models (London, Dec 17, 1997)

Finding Hidden Structure in Financial Time Series NBER Summer Institute (Cambridge, MA, Jul 16, 1997)

Data Mining in Finance IBM Research (Yorktown Heights, Jun 11, 1997)

Learning from Data in Finance and Business Leonard N. Stern School of Business, Affiliates Seminar (Mar 27, 1997)

Time Series Tools Computational Intelligence in Financial Engineering (New York, Mar 22, 1997)

Hidden Markov Experts RIKEN (Tokyo, Nov 1, 1996)

New Architectures for Time Series Analysis Neural Networks for Signal Processing (Keynote Lecture) (IEEE-NNSP, Kyoto, Sep 4, 1996)

Taking Time Seriously: The State of the State Department of Mathematical Engineering and Information Physics, University of Tokyo (Aug 29, 1996)

Neural Networks in Financial Engineering Monash University, Department of Business Systems (Melbourne, Jul 15, 1996)

Time Series and Chaos International Mathematical Society Meeting (IMS, Sydney, Jul 11, 1996)

Nonparametric Statistics: The Road Ahead Australian National University, Statistics Department (ANU, Canberra, Jul 7, 1996)

## 5. Other (Interactions, transitions, patent disclosures, etc.)

Due to my leaving CU and relinquishing the remainder of the first year of the grant, there unfortunately were no interactions, transitions, patent disclosures, or honors.

## 6. Personnel Supported

Andreas Weigend (PI) 3 months
Shanming Shi (graduate student) 2 1/2 months
Mark Choey (graduate student) 1 month
Mike Fellows (computer support) 1 month
Pat Libhart (secretary, 5% for 1 year)

A. Weigend, NYC Dec 10, 97